APPLICATION OF QUATERNIONS TO COMPUTATION WITH ROTATIONS

Working Paper, Stanford AI Lab, 1979[1]

by Eugene Salamin

Computer programs which operate on rotations may profitably represent these rotations by quaternions. The orthogonal matrix which performs a rotation by angle θ about axis $\mathbf{n}$ is derived. After developing some elementary properties of quaternions, it is shown how the same rotation can be represented by a unit quaternion. Formulas for interconversion between the two representations are derived. For computation with rotations, quaternions offer the advantage of requiring only 4 numbers of storage, compared with 9 numbers for orthogonal matrices. Composition of rotations requires 16 multiplications and 12 additions in quaternion representation, but 27 multiplications and 18 additions in matrix representation. Rotating a vector, with the rotation matrix in hand, requires 9 multiplications and 6 additions. However, if the matrix must be calculated from a quaternion, then this calculation needs 10 multiplications and 21 additions. The quaternion representation is more immune to accumulated computational error. A quaternion[2] Q which deviates from unicity can be fixed by $Q \leftarrow Q/|Q|$, however a matrix $\mathbf{R}$ which deviates from orthogonality must be fixed by the more involved calculation $\mathbf{R} \leftarrow \mathbf{R}(\mathbf{R}^T\mathbf{R})^{-1/2}$, or some approximation thereto. One can mentally picture the rotation group manifold as the 3-sphere with diametrically opposite points identified. The intrinsic metric of the manifold is the same as the metric resulting from the imbedding of the 3-sphere in Euclidean 4-space. This makes well-defined the concept of randomly chosen rotations. Finally, it is shown how Lorentz transformations in relativity can be implemented with complex quaternions.

1. Rotations and Orthogonal Matrices

Let a vector $\mathbf{x}$ be rotated by angle θ about the unit vector $\mathbf{n}$ to obtain the vector $\mathbf{x}'$. Write

$$\mathbf{x} = (\mathbf{x}\cdot\mathbf{n})\,\mathbf{n} + (\mathbf{x} - (\mathbf{x}\cdot\mathbf{n})\,\mathbf{n}).$$

The first term on the right is the component of $\mathbf{x}$ parallel to $\mathbf{n}$; it is unchanged by the rotation. The second term is the component of $\mathbf{x}$ perpendicular to $\mathbf{n}$. It is rotated into

$$(\mathbf{x} - (\mathbf{x}\cdot\mathbf{n})\,\mathbf{n})\cos\theta + \mathbf{n}\times(\mathbf{x} - (\mathbf{x}\cdot\mathbf{n})\,\mathbf{n})\sin\theta.$$

[1]Edited and TeX-formatted by Henry G. Baker (`hbaker1@pipeline.com`), Nov., 1995. Postings (`http://home.pipeline.com/~hbaker1/quaternion/stanfordaiwp79-salamin.ps.gz` and `.dvi.gz`) by permission of Eugene Salamin.

[2]Editor's note: to minimize confusion, we use *capital* letters for quaternions, *lower-case* letters for scalars, *bold capital* letters for matrices, and *bold lower-case* letters for vectors.

Collecting all terms, we see that

$$\mathbf{x}' = (\cos\theta)\,\mathbf{x} + (\sin\theta)\,\mathbf{n}\times\mathbf{x} + (1-\cos\theta)\,\mathbf{n}\,(\mathbf{n}\cdot\mathbf{x}). \tag{1.1}$$

We want to write (1.1) in matrix form

$$\mathbf{x}' = \mathbf{R}(\theta,\mathbf{n})\mathbf{x}, \tag{1.2}$$

where $\mathbf{R}(\theta,\mathbf{n})$ denotes the orthogonal matrix which performs a rotation by angle θ about axis $\mathbf{n}$. For this purpose, we set up a correspondence between vectors and antisymmetric matrices.

$$\mathbf{n} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} \qquad \leftrightarrow \qquad \mathbf{N} = \begin{pmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{pmatrix} \tag{1.3}$$

This correspondence is useful because $\mathbf{N}\mathbf{x} = \mathbf{n}\times\mathbf{x}$.

The characteristic equation of $\mathbf{N}$ is

$$\det(\lambda\mathbf{I} - \mathbf{N}) = \lambda^3 + (n_1^2 + n_2^2 + n_3^2)\,\lambda = 0.$$

Since $\mathbf{n}$ is a unit vector, the eigenvalues are distinct, and so the characteristic equation is also the minimal equation. Thus (Cayley-Hamilton theorem)

$$\mathbf{N}^3 = -\mathbf{N}. \tag{1.4}$$

From (1.1) and the definition of $\mathbf{N}$, the rotation matrix for rotation by a small angle ϵ is

$$\mathbf{R}(\epsilon,\mathbf{n}) = \mathbf{I} + \epsilon\mathbf{N} + O(\epsilon^2).$$

For a finite rotation,

$$\mathbf{R}(\theta,\mathbf{n}) = \mathbf{R}(\theta/k,\mathbf{n})^k = (\mathbf{I} + (\theta/k)\mathbf{N} + O(k^{-2}))^k.$$

Now let $k\to\infty$, and we get

$$\mathbf{R}(\theta,\mathbf{n}) = \exp(\theta\mathbf{N}) = e^{\theta\mathbf{N}}. \tag{1.5}$$

Expand $e^{\theta\mathbf{N}}$ into a power series. Use (1.4) to reduce all powers of $\mathbf{N}$ to $\mathbf{N}$ or $\mathbf{N}^2$. Finally, identify the power series coefficients of $\mathbf{N}$ and $\mathbf{N}^2$ with their trigonometric sums. The result is

$$e^{\theta\mathbf{N}} = \mathbf{I} + (\sin\theta)\,\mathbf{N} + (1-\cos\theta)\,\mathbf{N}^2. \tag{1.6}$$

By matrix multiplication,

$$\mathbf{N}^2 = \begin{pmatrix} -n_2^2 - n_3^2 & n_1n_2 & n_1n_3 \\ n_2n_1 & -n_1^2 - n_3^2 & n_2n_3 \\ n_3n_1 & n_3n_2 & -n_1^2 - n_2^2 \end{pmatrix} = \mathbf{nn} - \mathbf{I}, \tag{1.7}$$

in which $\mathbf{nn}$ denotes the dyadic product.

By substituting (1.7) into (1.6), and taking the product $\mathbf{x}' = \mathbf{R}(\theta,\mathbf{n})\mathbf{x}$, we find agreement with (1.1), which was derived directly from the definition of rotation.

2. Elementary Properties of Quaternions

Quaternions are elements of a certain 4-dimensional algebra, *i.e.* a vector space endowed with multiplication. For the moment, we make no assumption about the ground field of the vector space.

It is convenient to think of a quaternion, $Q = q_0 + \mathbf{q}$, as the sum of a scalar, q_0, and an ordinary 3-vector, $\mathbf{q}$. Then the quaternion product

$$R = PQ$$

is defined by

$$r_0 = p_0 q_0 - \mathbf{p}\cdot\mathbf{q}, \tag{2.1}$$

$$\mathbf{r} = p_0\,\mathbf{q} + \mathbf{p}\,q_0 + \mathbf{p}\times\mathbf{q}. \tag{2.2}$$

We will identify a scalar (field element) with a quaternion having vanishing vector part: $s = s + \mathbf{0}$. Likewise, we identify a 3-vector with a quaternion having vanishing scalar part: $\mathbf{v} = 0 + \mathbf{v}$. If $Q = q_0 + \mathbf{q}$, then by Q^* is meant the quaternion with the sign of its vector part changed, *i.e.* $Q^* = q_0 - \mathbf{q}$.

The quaternion algebra is also given a dot product.[3]

$$P\cdot Q = p_0 q_0 + \mathbf{p}\cdot\mathbf{q}. \tag{2.3}$$

A unit quaternion is a quaternion Q such that $Q\cdot Q = 1$. A null quaternion is a quaternion Q such that $Q\cdot Q = 0$. Non-zero null quaternions exist. Examples are $Q = (2,1,0,0)$ over Z_5, and $Q = (1,i,0,0)$ over the complex numbers. For this reason, the dot product is, in general, not an inner product.

The following properties are easily established, and are left as an exercise for the reader. Remember that the ground field is arbitrary in the current context.

1. The quaternion 1 is the multiplicative identity.

$$1Q = Q1 = Q.$$

2. Quaternion multiplication is associative.

$$P(QR) = (PQ)R = PQR.$$

Furthermore,[4]

$$\begin{aligned} PQR = p_0q_0r_0 &- p_0\,\mathbf{q}\cdot\mathbf{r} - q_0\,\mathbf{p}\cdot\mathbf{r} - r_0\,\mathbf{p}\cdot\mathbf{q} - [\mathbf{p},\mathbf{q},\mathbf{r}] \\ &+ q_0r_0\,\mathbf{p} + p_0r_0\,\mathbf{q} + p_0q_0\,\mathbf{r} \\ &+ p_0\,\mathbf{q}\times\mathbf{r} + q_0\,\mathbf{p}\times\mathbf{r} + r_0\,\mathbf{p}\times\mathbf{q} \\ &- (\mathbf{q}\cdot\mathbf{r})\,\mathbf{p} + (\mathbf{p}\cdot\mathbf{r})\,\mathbf{q} - (\mathbf{p}\cdot\mathbf{q})\,\mathbf{r}. \end{aligned} \tag{2.4}$$

[3] Editor's note: $P\cdot Q = \mathrm{Scalar}(P^*Q) = \mathrm{Scalar}(PQ^*) = \frac{PQ^*+(PQ^*)^*}{2}$.

[4] Editor's note: $[\mathbf{p},\mathbf{q},\mathbf{r}]$ is the "scalar triple product" — *i.e.*,

$$[\mathbf{p},\mathbf{q},\mathbf{r}] = \mathbf{p}\cdot(\mathbf{q}\times\mathbf{r}) = \begin{vmatrix} p_1 & p_2 & p_3 \\ q_1 & q_2 & q_3 \\ r_1 & r_2 & r_3 \end{vmatrix} = -\mathrm{Scalar}(\mathbf{pqr}) = -\frac{\mathbf{pqr} + (\mathbf{pqr})^*}{2}.$$

3. Quaternion multiplication is non-commutative. There exist P and Q such that $PQ \neq QP$. However, multiplication by a scalar is always commutative.

4. The * operator is an anti-homomorphism.

$$(PQ)^* = Q^* P^*.$$

5.

$$Q \cdot Q = QQ^* = Q^* Q.$$

6.

$$(PQ) \cdot (PQ) = (PQ) \cdot (QP) = (P \cdot P)(Q \cdot Q). \tag{2.5}$$

7.

$$(PQ) \cdot (PR) = (P \cdot P)(Q \cdot R). \tag{2.6}$$

3. Quaternion Description of Rotations

We now assume we are dealing with real quaternions. The dot product is now an inner product, and we can define the length of a quaternion Q as $|Q| = \sqrt{Q \cdot Q}$. It will be shown that the rotation matrices are in a 1-2 correspondence with the unit quaternions, *i.e.* those quaternions, Q, for which $|Q| = 1$.

Consider the transformation

$$X' = QXQ^*, \tag{3.1}$$

where Q is a unit quaternion. The scalar part, x_0, of X is transformed into $x_0' = (QQ^*)x_0 = x_0$. From (2.5), we see that $X' \cdot X' = X \cdot X$. Since the scalar part is left unchanged, we also have $\mathbf{x}' \cdot \mathbf{x}' = \mathbf{x} \cdot \mathbf{x}$, which is sufficient to prove that the transformation is a rotation. By specialization of (2.4), we see that the vector part, $\mathbf{x}$, of X is transformed into

$$\mathbf{x}' = (q_0^2 - \mathbf{q} \cdot \mathbf{q})\,\mathbf{x} + 2q_0\,\mathbf{q} \times \mathbf{x} + 2\mathbf{q}\,(\mathbf{q} \cdot \mathbf{x}). \tag{3.2}$$

Note the similarity in form between (1.1) and (3.2). It is reminiscent of the trigonometric double angle formulas. By choosing

$$Q = Q(\theta, \mathbf{n}) = \cos\frac{\theta}{2} + \sin\frac{\theta}{2}\,\mathbf{n},$$

(which is not much of a choice, since every unit quaternion is of this form) we make the two equations identical.

This can also be written in a form similar to (1.5). First, note that, *as a quaternion product*, $\mathbf{n}^2 = -1$. By manipulation of the power series, just as was done for rotation matrices, we find

$$Q(\theta, \mathbf{n}) = e^{(\theta/2)\mathbf{n}} = \cos\frac{\theta}{2} + \sin\frac{\theta}{2}\,\mathbf{n}. \tag{3.3}$$

We have now exhibited the correspondence between unit quaternions and proper orthogonal matrices as representations of rotations.

$$Q(\theta, \mathbf{n}) \leftrightarrow \mathbf{R}(\theta, \mathbf{n}) \tag{3.4}$$

This correspondence is a group homomorphism because $P(QXQ^*)P^* = PQXQ^*P^* = (PQ)X(PQ)^*$. It is not quite an isomorphism, since both $Q(\theta, \mathbf{n})$ and $-Q(\theta, \mathbf{n})$ correspond to the same orthogonal matrix.

Given the elements of a unit quaternion, Q, we can use (3.2) to obtain the elements of the corresponding rotation matrix, $\mathbf{R}$.

$$\mathbf{R} = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(-q_0q_3 + q_1q_2) & 2(q_0q_2 + q_1q_3) \\ 2(q_0q_3 + q_2q_1) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(-q_0q_1 + q_2q_3) \\ 2(-q_0q_2 + q_3q_1) & 2(q_0q_1 + q_3q_2) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix}. \tag{3.5}$$

Conversely, given the elements of a proper orthogonal matrix, $\mathbf{R}$, we want to compute the elements of a corresponding unit quaternion, Q. From the diagonal elements of (3.5) we get

$$\begin{aligned} q_0^2 &= \tfrac{1}{4}(1 + \mathbf{R}_{11} + \mathbf{R}_{22} + \mathbf{R}_{33}), \\ q_1^2 &= \tfrac{1}{4}(1 + \mathbf{R}_{11} - \mathbf{R}_{22} - \mathbf{R}_{33}), \\ q_2^2 &= \tfrac{1}{4}(1 - \mathbf{R}_{11} + \mathbf{R}_{22} - \mathbf{R}_{33}), \\ q_3^2 &= \tfrac{1}{4}(1 - \mathbf{R}_{11} - \mathbf{R}_{22} + \mathbf{R}_{33}). \end{aligned} \tag{3.6}$$

From the off-diagonal elements of (3.5) we get

$$\begin{aligned} q_0q_1 &= \tfrac{1}{4}(\mathbf{R}_{32} - \mathbf{R}_{23}), \\ q_0q_2 &= \tfrac{1}{4}(\mathbf{R}_{13} - \mathbf{R}_{31}), \\ q_0q_3 &= \tfrac{1}{4}(\mathbf{R}_{21} - \mathbf{R}_{12}), \\ q_1q_2 &= \tfrac{1}{4}(\mathbf{R}_{12} + \mathbf{R}_{21}), \\ q_1q_3 &= \tfrac{1}{4}(\mathbf{R}_{13} + \mathbf{R}_{31}), \\ q_2q_3 &= \tfrac{1}{4}(\mathbf{R}_{23} + \mathbf{R}_{32}). \end{aligned} \tag{3.7}$$

Choose the equation from (3.6) which yields the q_k of largest magnitude. In particular, if we find some $q_k^2 \geq 1/4$, then no other element of Q can be larger. Since both Q and $-Q$ are solutions to the same $\mathbf{R}$, we can take either sign of the square root. Once we have solved for one q_k, the three equations in (3.7) which contain q_k on the left hand side can be solved for the remaining elements of Q.

4. Comparison Between Quaternions and Orthogonal Matrices

If rotations are represented by 3×3 orthogonal matrices, then it is necessary to store all 9 elements. Composing rotations requires multiplying the two matrices. For each product matrix element, we must do 3 multiplications and 2 additions, for a total of 27 multiplications and 18 additions.

If rotations are represented by quaternions, then only 4 elements need to be stored. Composition of rotations corresponds to quaternion multiplication. When the quaternion product, (2.1) and (2.2) is written out in terms of actual elements, we see that computing each product element requires 4 multiplications and 3 additions, for a total of 16 multiplications and 12 additions.

When rotating a vector, if we already possess the rotation matrix, then only 9 multiplications and 6 additions are needed. If we have to compute the rotation matrix from the quaternion, then we can follow the procedure of (3.5). First compute all the quadratic monomials $q_j q_k$, which takes 10 multiplications. Next, each of the three diagonal terms can be calculated with 3 additions. Finally, each of the six off-diagonal terms can be calculated with 2 additions, provided multiplication by 2 is considered as an addition. The total computation in obtaining the matrix is 10 multiplications and 21 additions.

Buildup of computational error will cause the quaternion to become of non-unit length. This can be fixed by dividing the quaternion by $|Q|$. In contrast, fixing a non-orthogonal matrix is much more complicated (see Appendix A). If $\mathbf{A}$ is an almost orthogonal matrix, then $\mathbf{A}$ can be re-orthogonalized by replacing it by $\mathbf{R} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1/2}$.

But we can even avoid having to rescale the quaternion provided we rotate vectors according to

$$\mathbf{x}' = \frac{Q\mathbf{x}Q^*}{Q \cdot Q}.$$

No such option exists when orthogonal matrices are used.

5. The Rotation Group Manifold

We already know what the rotation group manifold looks like topologically. It is a 3-sphere, *i.e.* the 3-dimensional surface of a 4-dimensional ball, with diametrically opposite points identified, since Q and $-Q$ represent the same rotation. This space is usually called a projective 3-space, and denoted by $\mathcal{P}^3$. Equivalently, the points of $\mathcal{P}^3$ can be taken to be the lines through the origin in Euclidean 4-space.

The group manifold inherits a metric, the dot product (2.3), by virtue of its imbedding in Euclidean space. A group manifold also has an intrinsic metric, which is defined by the requirement that it be invariant under group multiplication. If P, Q, and R are group elements, *i.e.* unit quaternions, then the invariant metric must satisfy $(PQ) \cdot (PR) = (QP) \cdot (RP) = Q \cdot R$. But, by (2.6), the inherited metric already satisfies this condition. This validates in total the picture of the rotation group manifold as a 3-sphere (with opposite points identified).

The invariant metric implies a measure on the group manifold, and thus allows probability distributions to be defined. This is useful for Monte Carlo simulations where we need to choose "random rotations". Random rotations are to be chosen uniformly distributed over the 3-sphere.

In order to carry out integrations over the group manifold, we can use hyperspherical coordinates.

$$\begin{aligned} q_0 &= \cos\alpha, \\ q_1 &= \sin\alpha\,\cos\beta, \\ q_2 &= \sin\alpha\,\sin\beta\,\cos\gamma, \\ q_3 &= \sin\alpha\,\sin\beta\,\sin\gamma. \end{aligned} \tag{5.1}$$

To cover the $\mathcal{P}^3$, we only need the region $q_0 \geq 0$. The range of the angles is

$$\begin{aligned} &0 \leq \alpha \leq \pi/2, \\ &0 \leq \beta \leq \pi, \\ &0 \leq \gamma \leq 2\pi. \end{aligned}$$

The length element is

$$ds^2 = \, d\alpha^2 + \sin^2\!\alpha\, d\beta^2 + \sin^2\!\alpha\, \sin^2\!\beta\, d\gamma^2, \tag{5.2}$$

and the volume element is

$$d\Omega = \sin^2\!\alpha\, \sin\beta\, d\alpha\, d\beta\, d\gamma. \tag{5.3}$$

The total group volume is

$$\Omega = \int_0^{\pi/2} d\alpha \int_0^{\pi} d\beta \int_0^{2\pi} d\gamma\, \sin^2\!\alpha\, \sin\beta = \pi^2. \tag{5.4}$$

As an application, let's calculate the mean value of the rotation angle, θ, of randomly chosen rotations. Since $q_0 = \cos(\theta/2) = \cos\alpha$, we have

$$\langle\theta\rangle = 2\langle\alpha\rangle = \frac{1}{\Omega}\int 2\alpha\, d\Omega = \frac{8}{\pi}\int_0^{\pi/2} \alpha \sin^2\!\alpha\, d\alpha = \frac{\pi}{2} + \frac{2}{\pi} = 126.5\,\text{deg}.$$

Appendix A. Re-orthogonaliztion of Matrices

Let $\mathbf{A}$ be a matrix which is supposed to be orthogonal, but isn't, because of accumulated computational error. We want to replace $\mathbf{A}$ by the "nearest" orthogonal matrix. Let $\mathbf{X}$ be the unknown replacement matrix. $\mathbf{X}$ can be found by minimizing the quantity

$$f = \sum_{i,j} \left((\mathbf{X}_{ij} - \mathbf{A}_{ij})^2 - \mathbf{M}_{ij} (\mathbf{X}^T\mathbf{X} - \mathbf{I})_{ij} \right).$$

The first term on the right hand side assures that $\mathbf{X}$ is close to $\mathbf{A}$. In the second term, $\mathbf{M}$ is a Lagrange multiplier matrix. We minimize f by variation of $\mathbf{X}$ and $\mathbf{M}$. Differentiation of f with respect to the elements of $\mathbf{M}$ gives the condition, $\mathbf{X}^T\mathbf{X} = \mathbf{I}$, that $\mathbf{X}$ be orthogonal. Since $\mathbf{X}^T\mathbf{X} - \mathbf{I}$ is always symmetric, regardless of $\mathbf{X}$, there are only 6 constraints, not 9. So we need only 6 independent Lagrange multipliers, *i.e.* we assume that $\mathbf{M}$ is symmetric.

Then, differentiating f with respect to the elements of $\mathbf{X}$ results in

$$\mathbf{X} - \mathbf{A} - \mathbf{X}\mathbf{M} = \mathbf{0}. \tag{A.1}$$

The solution of (A.1) proceeds as follows.

$$\begin{aligned}
\mathbf{X} &= \mathbf{A}(\mathbf{I} - \mathbf{M})^{-1} \\
\mathbf{X}^T &= (\mathbf{I} - \mathbf{M})^{-1}\mathbf{A}^T \\
\mathbf{X}^T\mathbf{X} &= (\mathbf{I} - \mathbf{M})^{-1}\mathbf{A}^T\mathbf{A}(\mathbf{I} - \mathbf{M})^{-1} = \mathbf{I} \\
\mathbf{A}^T\mathbf{A} &= (\mathbf{I} - \mathbf{M})^2 \\
\mathbf{I} - \mathbf{M} &= (\mathbf{A}^T\mathbf{A})^{1/2} \\
\mathbf{X} &= \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1/2}
\end{aligned}$$

Note that since $\mathbf{A}$ is almost orthogonal, $\mathbf{A}^T\mathbf{A}$ is almost the identity, and so exactly one of its square roots is also almost the identity. This is obviously the square root to be used.

Appendix B. Quaternions and Relativity

This appendix will mainly be of interest to the reader acquainted with the special theory of relativity. Quaternions are now complex, and the dot product defined in (2.3) and the definition of unit quaternion are carried over exactly as stated, *i.e.* there is no complex conjugation. When we need complex conjugation, we denote it by an overline.

We will consider the following two transformations.

$$U' = QU\overline{Q}^*, \tag{B.1}$$

$$U' = QUQ^*, \tag{B.2}$$

in which Q is a complex unit quaternion. (B.1) is the Lorentz transformation of 4-vectors, e.g. space-time displacements, or momentum-energy. (B.2) is the Lorentz transformation of the electromagnetic field.

Starting with (B.1), we see from (2.5) that $U' \cdot U' = U \cdot U$. Suppose the vector part of U is real and the scalar part of U is imaginary, *i.e.* we have $U = it + \mathbf{x}$, where t and $\mathbf{x}$ are real. Also let $U' = it' + \mathbf{x}'$. Then the invariance of the dot product is reformulated as

$$-t'^2 + \mathbf{x}' \cdot \mathbf{x}' = -t^2 + \mathbf{x} \cdot \mathbf{x}, \tag{B.3}$$

which is the familiar condition defining the Lorentz transformation.

We proceed to prove that (B.1) is a Lorentz transformation. We must show that t' and $\mathbf{x}'$ are real, for otherwise (B.3) would not be correctly interpreted. Note that t and $\mathbf{x}$ are real if and only if $U = it + \mathbf{x}$ satisfies $\overline{U} = -U^*$, and likewise for U'. So, to prove that (B.1) is a Lorentz transformation, we must show that $\overline{U} = -U^*$ implies $\overline{U}' = -U'^*$.

$$\overline{U}' = \overline{QU}Q^* = -\overline{Q}U^*Q^* = -\left(QU\overline{Q}^*\right)^* = -U'^*.$$

As Q varies over all unit quaternions, the (B.1) transformations can, at most, generate the proper Lorentz transformations, *i.e.* those which involve no space or time reflections. To show that all the proper Lorentz transformations are obtained, we only need to obtain a generating set. Such a generating set consists of the rotations and the Lorentz boosts, or pure velocity transformations.

The rotations are obtained by choosing Q real. Then Q necessarily has the canonical form (3.3) with θ and $\mathbf{n}$ real, and we get

$$Q\,(it+\mathbf{x})\,\overline{Q}^{*} = Q\,(it+\mathbf{x})\,Q^{*} = it + Q\mathbf{x}Q^{*}.$$

The time part is unchanged, and the space part is rotated, exactly as in section 3.

The pure velocity boosts are obtained by choosing Q to have a real scalar part and imaginary vector part. Then Q necessarily has the canonical form

$$Q = \cosh\frac{\alpha}{2} + i\sinh\frac{\alpha}{2}\,\mathbf{m}, \tag{B.4}$$

where α is real and $\mathbf{m}$ is a real unit vector. We find, by using (2.4),

$$\begin{aligned} U' &= QU\overline{Q}^{*} = QUQ \\ &= (q_0^2 - \mathbf{q}\cdot\mathbf{q})u_0 - 2q_0\,\mathbf{q}\cdot\mathbf{u} + \mathbf{u} - 2\mathbf{q}\,(\mathbf{q}\cdot\mathbf{u}) + 2q_0\,\mathbf{q}\,u_0. \end{aligned} \tag{B.5}$$

Substitute (B.4) into (B.5), let $U = it + \mathbf{x}$, apply the hyperbolic double angle formulas, and finally, separate the imaginary scalar part from the real vector part.

$$\begin{aligned} t' &= (\cosh\alpha)\,t - (\sinh\alpha)\,\mathbf{m}\cdot\mathbf{x} \\ \mathbf{x}' &= (\cosh\alpha)\,\mathbf{mm}\cdot\mathbf{x} + (1-\mathbf{mm})\cdot\mathbf{x} - (\sinh\alpha)\,\mathbf{m}t \end{aligned} \tag{B.6}$$

Let's put (B.6) into its more familiar form. Let $\mathbf{m}$ be along the z-axis, let the velocity be $v/c = \tanh\alpha$, and note that $\cosh\alpha = 1/\sqrt{1-v^2/c^2} = \gamma$. Then (B.6) becomes

$$\begin{aligned} t' &= \gamma(t - (v/c)z), \\ x' &= x, \\ y' &= y, \\ z' &= \gamma(z - (v/c)t). \end{aligned}$$

In general, (B.6) is the Lorentz transformation for velocity $v/c = (\tanh\alpha)\mathbf{m}$.

Next, we examine the transformation (B.2). Note that (B.2) is algebraically identical to (3.1), except that now the quaternions are all complex. Just as in section 3, $U'\cdot U' = U\cdot U$, and $u_0' = u_0$, so $\mathbf{u}'\cdot\mathbf{u}' = \mathbf{u}\cdot\mathbf{u}$. Anticipating the type of tensor whose transformation is effected by (B.2), we let $\mathbf{u} = \mathbf{E} + i\mathbf{B}$, where $\mathbf{E}$ and $\mathbf{B}$ are real.[5] Then we see that $\mathbf{E}\cdot\mathbf{E} - \mathbf{B}\cdot\mathbf{B}$ and $\mathbf{E}\cdot\mathbf{B}$ are invariants of the transformation.

[5] Editor's note: we break our convention here and allow bold capitals to denote the familiar electric and magnetic vectors.

When Q is a pure rotation, and has canonical form (3.3), $\mathbf{E}$ and $\mathbf{B}$ are independently rotated. When Q is a pure velocity boost, substitute the canonical form (B.4) into (3.2) and apply the double angle formulas to get

$$\mathbf{u}' = \mathbf{mm} \cdot \mathbf{u} + (\cosh\alpha)\,(1 - \mathbf{mm}) \cdot \mathbf{u} + 2i(\sinh\alpha)\,\mathbf{m} \times \mathbf{u}.$$

With separation into real and imaginary parts, this becomes

$$\begin{aligned} \mathbf{E}' &= \mathbf{mm} \cdot \mathbf{E} + (\cosh\alpha)\,(1 - \mathbf{mm}) \cdot \mathbf{E} - (\sinh\alpha)\,\mathbf{m} \times \mathbf{B}, \\ \mathbf{B}' &= \mathbf{mm} \cdot \mathbf{b} + (\cosh\alpha)\,(1 - \mathbf{mm}) \cdot \mathbf{B} + (\sinh\alpha)\,\mathbf{m} \times \mathbf{E}. \end{aligned} \tag{$B.7$}$$

To put (B.7) into the familiar form of the Lorentz transformation of the electric and magnetic fields, again choose $\mathbf{m}$ to lie along the z-axis.

$$\begin{aligned} \mathbf{E}'_x &= \gamma(\mathbf{E}_x + (v/c)\mathbf{B}_y), \\ \mathbf{B}'_x &= \gamma(\mathbf{B}_x - (v/c)\mathbf{E}_y), \\ \mathbf{E}'_y &= \gamma(\mathbf{E}_y - (v/c)\mathbf{B}_x), \\ \mathbf{B}'_y &= \gamma(\mathbf{B}_y + (v/c)\mathbf{E}_x), \\ \mathbf{E}'_z &= \mathbf{E}_z, \\ \mathbf{B}'_z &= \mathbf{B}_z. \end{aligned}$$